

Decoupled Software Pipelining: A Promising Technique to Exploit Thread-Level Parallelism

Guilherme Ottoni

with
Ram Rangan, Neil Vachharajani,
David I. August

Liberty Research Group
Department of Computer Science
Princeton University

<http://www.liberty-research.org>

A Fundamental Change...

Clock rate limited by:

- Power delivery
- Heat dissipation

Last year: P4 4.0 GHz...

Source: Intel, Wikipedia, Sutter/Dr. Dobbs Journal

The Liberty Research Group

The Response: CMP

For:

- legacy applications
- single threaded
- sequential codes
- C/C++

Speedup over single core:

0.0%

Need To Extract Parallelism!

IBM Power 5 (1.9GHz) Die Photo: Source IBM

The Liberty Research Group

Typical Piece of Sequential Code

```
while(ptr = ptr->next) {
    ptr->val = ptr->val + 1;
}
```

```

A: r1 = M[r1]
B: r2 = r1 + 4
C: r3 = M[r2]
D: r4 = r3 + 1
E: M[r2] = r4
F: p1 = r1 != 0
G: br p1, Loop
    
```

ILP works well, but:

- No way to create concurrent threads!
- Can be trapped by variable latency instructions.

The Liberty Research Group

How do we create concurrent threads?

Key: Pipelined Parallelism: Dependent and Decoupled Threads!

Inspired by:
Architectures: CMP, [Smith ISCA-82]
Languages: StreamIT, Stampede, ...
Special-purpose compilers:
Daecomp, [Du SC-03, Yang IPDPS-03]

The Liberty Research Group

Decoupled Software Pipelining (DSWP)

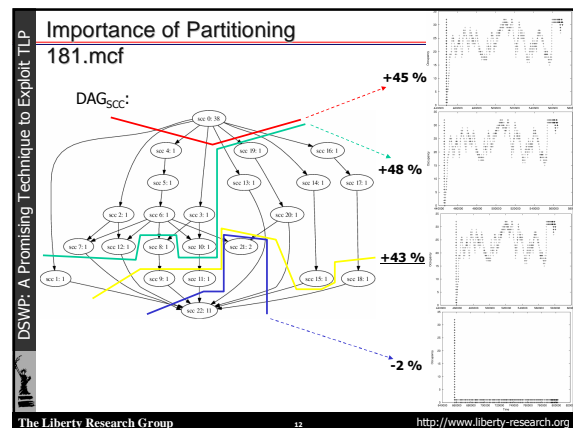
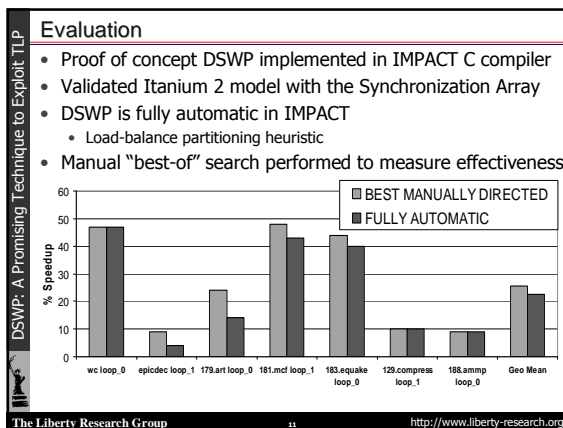
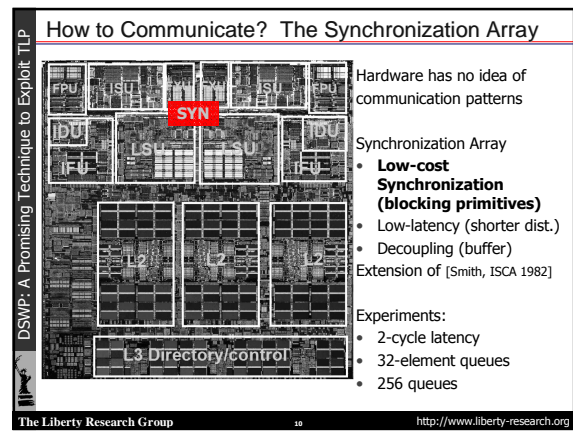
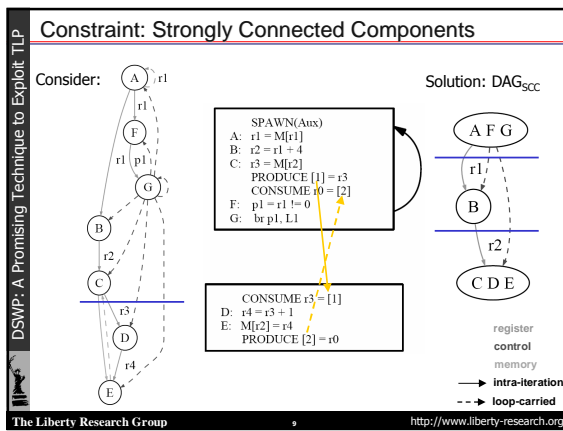
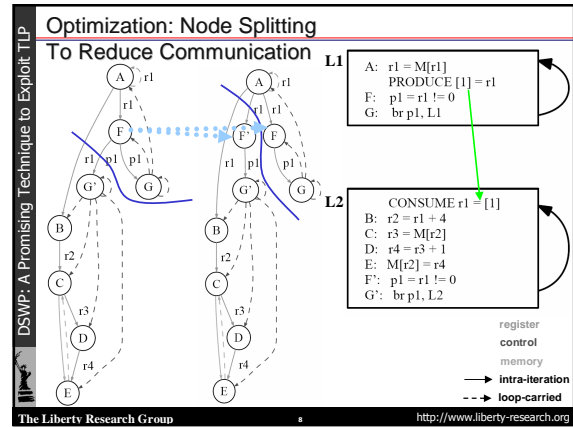
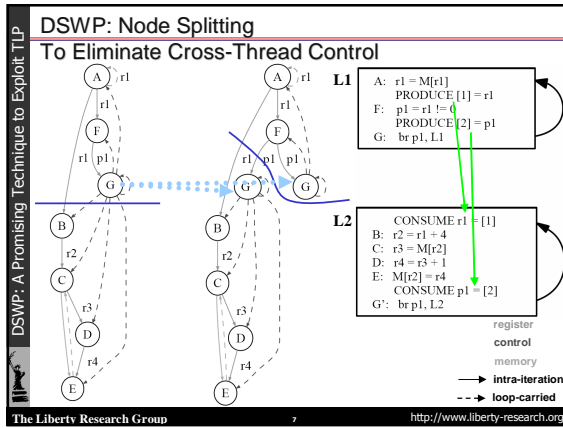
```

A: r1 = M[r1]
B: r2 = r1 + 4
C: r3 = M[r2]
D: r4 = r3 + 1
E: M[r2] = r4
F: p1 = r1 != 0
G: br p1, Loop
    
```

Dependence Graph

Legend:
— intra-iteration
--- loop-carried

The Liberty Research Group



DSWP: A Promising Technique to Exploit TLP

Conclusion

Chip multiprocessors are inevitable due to fundamental problems

Automatically threading sequential programs is possible!

- Pipelined Parallelism: Create Dependent Threads
- Techniques to handle control flow (inner loops not enough)
- Synchronization Array necessary for efficient communication
- Decoupling for latency tolerance

ILP research must be extended/recast:

- Analysis: Memory alias, shape
- Optimizations: Node splitting, accumulator expansion
- Scope: Whole program optimization

The Liberty Research Group19http://www.liberty-research.org